DOCUMENT RESUME

ED 433 353 TM 030 022

AUTHOR	Aaron, Bruce; Kromrey, Jeffrey D.; Ferron, John
TITLE	Equating "r"-based and "d"-based Effect Size Indices: Problems with a Commonly Recommended Formula.
PUB DATE	1998-11-03
NOTE	13p.; Paper presented at the Annual Meeting of the Florida Educational Research Association (43rd, Orlando, FL, November 2-4, 1998).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Correlation; *Effect Size; *Estimation (Mathematics); *Meta Analysis; Sample Size

ABSTRACT

Two general categories comprise the various effect size indices that have been proposed for use in meta-analysis: (1) the "d"-type estimator (based on magnitude of mean difference); and (2) the "r"-type estimator (based on magnitude of correlation). In meta-analyses, researchers often must convert these effect size indices to a common metric to aggregate and synthesize results from various studies empirically. A commonly recommended formula for equating mean-difference effect sizes and correlational effect sizes was found to lead to inaccurate results, particularly with small sample sizes. A correct formula for converting "d"-based and "r"-based effect size indices is presented. Results of applying the common and corrected formulas are illustrated for a variety of data conditions at various sample sizes and effect sizes, suggesting that bias as large as 20 percent results from the common formula, something that can be avoided by applying the alternative equation. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 433 353

Equating *r*-based and *d*-based effect size indices: Problems with a commonly recommended formula

Bruce Aaron
Jeffrey D. Kromrey
John Ferron

University of South Florida

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Bruce Aaron
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

TM030022

A paper presented at the annual meeting of the
Florida Educational Research Association, Orlando, Florida, November 3, 1998

BEST COPY AVAILABLE

Abstract

Two general categories comprise the various effect size indices that have been proposed for use in meta-analysis: (a) the *d*-type estimator (based on magnitude of mean difference), and (b) the *r*-type estimator (based on magnitude of correlation). In meta-analyses, researchers often must convert these effect size indices to a common metric to empirically aggregate and synthesize results from various studies. A commonly recommended formula for equating mean-difference effect sizes and correlational effect sizes was found to lead to inaccurate results, particularly with small sample sizes. A correct formula for converting *d*-based and *r*-based effect size indices is presented. Results of applying the common and corrected formulas are illustrated for a variety of data conditions at various sample sizes and effect sizes, suggesting that bias as large as 20% results from the common formula, all of which can be avoided by applying the alternative equation.

Equating *r*-based and *d*-based effect size indices: Problems with a commonly recommended formula

Purpose

The purpose of this paper is to inform researchers about inaccuracies that can result from applying a commonly cited formula for converting effect size indices based on mean differences to effect size indices based on correlation. A corrected formula for equating these *d*-based and *r*-based effect size indices is presented. In addition, a demonstration of the potential bias of the commonly recommended formula is provided under data conditions displaying a variety of effect sizes, sample sizes, and degrees of inequality between sample sizes.

Background

Essential to synthesis of research are the processes of comparing and combining the results of studies pertaining to a particular research question. Meta-analysis is a prominent method of synthesis that uses statistical procedures to summarize findings from various individual studies, seeking an objective integration of the results of research conducted in an area of particular interest. In recent years, meta-analytic methods for summarizing research results across studies have proliferated in the social, behavioral, and biomedical sciences (Wolf, 1986).

Generally, meta-analysis summarizes independent statistical tests of a common hypothesis, and the strength of the relationship of interest. The latter is summarized by calculating effect sizes. Various effect size measures have been proposed for different test statistics (Rosenthal, 1994). Two general categories of effect size are conceptualized as (a) the *d*-type estimator (based on magnitude of mean difference), and (b) the *r*-type estimator (based on magnitude of correlation). The *d*-type effect size estimators derive a standardized mean difference that Cohen (1988) describes as a "pure number, one free of our original measurement unit with which to index what can be alternatively called the degree of departure from the null hypothesis of the alternative

hypotheses . . . ". This difference between two sample means, expressed in standard deviation units, is frequently referred to as Cohen's d, and is represented as:

$$d = \frac{\overline{X}_1 - \overline{X}_2}{\hat{\sigma}_{pooled}} \tag{1}$$

In correlational studies, an effect size can be described as a signal-to-noise ratio, and estimates the strength of relationship between interval or ratio scale measures. Such an effect size estimate is offered by Cohen (1983, p. 117) as:

$$f^2 = \frac{r^2}{1 - r^2} \tag{2}$$

In meta-analytic reviews of literature, the set of studies defined for review by the researcher often comprises diverse analysis procedures, requiring use of both d- and r-type estimators. In order to aggregate and synthesize these results, one must translate the various effect size estimates into a common metric. The statistic usually chosen as the basis for this common metric is r (the Pearson Product-Moment coefficient) (Wolf, 1986). Procedures for converting d to r have been discussed by Cohen (1965, 1988), Friedman (1968), Glass, McGraw, and Smith (1981), Rosenthal (1984), and Wolf (1986), with the following formula provided for expressing d as r:

$$r = \sqrt{\frac{d^2}{d^2 + 4}} \tag{3}$$

Equivalently, a conversion from r to d is offered as:

$$d = \frac{2r}{\sqrt{1 - r^2}} \tag{4}$$

These formulas, however, negatively bias estimates of *r* relative to *d*, particularly when sample sizes are unequal. This was accounted for by Hedges and Olkin (1985, p. 77), who offered the conversion from *d* to *r* appropriately as an approximation:

$$r^2 = \frac{d^2}{d^2 + (n_1 + n_2 - 2)/\tilde{n}} \tag{5}$$

where $\tilde{n} = (n_1 n_2)/(n_1 + n_2)$. When $n_1 = n_2 = n$ formula 5 reduces to:

$$r^2 = \frac{d^2}{d^2 + 4(n-1)/n} \cong \frac{d^2}{d^2 + 4} \tag{6}$$

This indicates correctly both the approximation provided by formula 3, and the conditional appropriateness of formula 3 for conversion when sample sizes are equal.

One way of illustrating the relationship of *d*- and *r*- based estimators is by considering the test of significance for differences between two sample group means, where the *t*-test of the standardized mean differences between groups, and the *t*-test associated with the point biserial correlation between group and dependent variable, are identical. An exact equation for the conversion of *d* to *r* can be derived from the relationship of the test of significance and effect size for a given study, noted by Rosenthal (1994) as

Test Statistic = Effect Size × Size of Study,

or, for the specific case in which effect size is expressed as the standardized mean difference between two groups,

$$t = \frac{\overline{X_1} - \overline{X_2}}{S_p} \times \frac{1}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\overline{X_1} - \overline{X_2}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{7a}$$

It follows that

$$t = \frac{\overline{X_1} - \overline{X_2}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = r\sqrt{\frac{N-2}{1-r^2}} \tag{7b}$$

$$d = r\sqrt{\frac{N-2}{1-r^2}}\left(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) \tag{7c}$$

$$d^2 = r^2\left(\frac{N-2}{1-r^2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \tag{7d}$$

$$\left(1-r^2\right)d^2 = r^2(N-2)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \tag{7e}$$

$$d^2 - d^2r^2 = r^2\left[(N-2)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right] \tag{7f}$$

$$d^2 = r^2\left[(N-2)\left(\frac{1}{n_1} + \frac{1}{n_2}\right) + d^2\right] \tag{7g}$$

$$r^2 = \frac{d^2}{d^2 + (N-2)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \tag{7h}$$

$$= \frac{d^2}{d^2 + \frac{N^2 - 2N}{n_1n_2}} \tag{8}$$

Equation 8 can be seen as an alternative derivation of a conversion formula offered by Cohen (1988, p. 24) for the general case (i.e., applicable to unequal sample sizes), with p as the proportion in one group on a dichotomous variable, and $q = 1 - p$:

$$r^2 = \frac{d^2}{d^2 + \left(\frac{1}{pq}\right)} \tag{9}$$

$$= \frac{d^2}{d^2 + \frac{N^2}{n_1n_2}} \tag{9a}$$

Therefore, corrected equation 8 can be seen to differ from equation 9a by replacement of N^2 with $N^2 - 2N$ in the second term of the denominator.

For balanced situations, in which $n_1 = n_2$, corrected equation 7h reduces as follows:

$$r^2 = \frac{d^2}{d^2 + (N-2)\left(\frac{4}{N}\right)} \tag{10}$$

$$= \frac{d^2}{d^2 + 4 - \left(\frac{8}{N}\right)} \tag{10a}$$

$$r = \sqrt{\frac{d^2}{d^2 + 4 - \left(\frac{8}{N}\right)}} \tag{10b}$$

Thus, for balanced situations, corrected equation 10b is seen to differ from commonly recommended formula 3 by subtraction of the last term in the denominator.

Data Source and Method

To provide a demonstration of the bias associated with formulas 3 and 9, the equations were applied to data conditions for pairs of samples having varying effect sizes, sample sizes, and degrees of inequality between sample sizes. These conditions represented sample pairs with standardized mean differences (d - type effect sizes) of 0.2, 0.5, and 0.8. These levels represent treatment effects ranging from small to large, as suggested by Cohen (1988). Sample sizes for cases in which $n_1 = n_2 = n$ were 5, 10, 20, 30, 40, 50, and 100. For the unbalanced sample size comparisons, $n_1 + n_2 = 20$, and the magnitude of inequality between sample sizes ranged from zero (where $n_1 = n_2 = 10$), to a ratio of 9:1 (where $n_1 = 18$ and $n_2 = 2$). These conditions were deemed to provide a practical range for investigating the differences resulting from application of the commonly recommended formulas for equating d-type and r-type effect size estimators and application of the alternative corrected formulas.

Values of r derived by the commonly recommended formulas (formulas 3 and 9) and the values of r provided by the corrected formulas (formulas 8 and 10b) were calculated for each unique set of conditions. These values of r were then used to calculate f^2. Bias is shown as the difference between the tabled values of r and f^2 derived by the common and actual formulas. The proportion of the actual effect size estimate comprising this bias is shown as proportional bias.

Results and Conclusions

Results presented in tables 1 and 2 demonstrate the magnitude of discrepancy found when applying the commonly recommended formula, and the alternative equations shown as formulas 8 and 10, for balanced and unbalanced sample sizes. Differences between the tabled values of r and f^2 derived by the commonly recommended formulas, and the values of these effect size estimates as provided by the corrected formulas, are shown as bias, with proportional bias indicating the size of this bias as a proportion of the actual effect size estimate (calculated with the corrected formula).

As shown in tables 1 and 2, bias associated with the common formula is particularly problematic for small sample sizes and when converting a d-based effect size estimate to an f^2-based effect size estimate. Among the conditions examined, the largest bias (20%) was found when translating d to f^2 for $n_1 = n_2 = 5$. Even for $n_1 = n_2 = 100$, a bias of 1% was avoided with the corrected formula. With unbalanced samples, conversion of d to f^2 incurred a bias of 10%, and estimation of r from d biased effect size estimates by 4% to 5%. For unbalanced samples, the extent of inequality in sample sizes had no practical effect on proportional bias. When estimating f^2 from d for these conditions, this bias was constant at 10%. The size of treatment effect demonstrated no practical influence on the extent of bias found for estimates of either f^2 or r, given any particular combination of n_1 and n_2. Given these results, use of the corrected formulas by analysts seeking to synthesize d-type and r-type estimates in meta-analysis is recommended.

Limitations

The inaccuracy resulting from application of the more commonly cited, approximate equation will differ according to the range of n, and might be negligible for certain conditions. For example, less bias was found for larger samples and when estimating r from d. Discrepancies tended to be negligible, therefore, when estimating r from d with balanced sample sizes of 50 or more.

Implications

In meta-analyses, researchers often encounter studies utilizing diverse research designs in which some effect sizes might be based on standardized differences in means, while others might be based on the correlation between variables. In order to empirically aggregate and synthesize the results from such diverse analyses, the meta-analyst must convert these effect sizes to a common metric. The commonly recommended formulas for equating correlational effect sizes and mean-difference effect sizes were found to lead to inaccurate results, particularly for small samples and when translating d to f^2. Corrected formulas for converting d-based and r-based effect size indices were presented and demonstrated. Since the corrected formulas are straightforward, and the accuracy of meta-analysis can be improved by their use, this is an important detail for researchers concerned with meta-analytic methods.

Table 1. Estimations of r and f^2 with equal sample sizes

N1	N2	*d*	Common Formula	Actual (Corrected Formula)	Bias *	Proportional Bias **
			Estimating *r* from *d*			
5	5	0.2	.0995	.1111	- .0116	- .1045
10	10	0.2	.0995	.1048	- .0053	- .0508
20	20	0.2	.0995	.1021	- .0026	- .0251
30	30	0.2	.0995	.1012	- .0017	- .0166
40	40	0.2	.0995	.1008	- .0013	- .0125
50	50	0.2	.0995	.1005	- .0010	- .0100
100	100	0.2	.0995	.1000	- .0005	- .0050
5	5	0.5	.2425	.2692	- .0267	- .0990
10	10	0.5	.2425	.2548	- .0123	- .0482
20	20	0.5	.2425	.2485	- .0059	- .0238
30	30	0.5	.2425	.2464	- .0039	- .0158
40	40	0.5	.2425	.2454	- .0029	- .0118
50	50	0.5	.2425	.2449	- .0023	- .0095
100	100	0.5	.2425	.2437	- .0011	- .0047
5	5	0.8	.3714	.4083	- .0368	- .0903
10	10	0.8	.3714	.3885	- .0171	- .0441
20	20	0.8	.3714	.3797	- .0083	- .0218
30	30	0.8	.3714	.3768	- .0055	- .0145
40	40	0.8	.3714	.3755	- .0041	- .0108
50	50	0.8	.3714	.3746	- .0032	- .0087
100	100	0.8	.3714	.3730	- .0016	- .0043
			Estimating f^2 from *d*			
5	5	0.2	.0100	.0125	- .0025	- .2000
10	10	0.2	.0100	.0111	- .0011	- .1000
20	20	0.2	.0100	.0105	- .0005	- .0500
30	30	0.2	.0100	.0103	- .0003	- .0333
40	40	0.2	.0100	.0103	- .0003	- .0250
50	50	0.2	.0100	.0102	- .0002	- .0200
100	100	0.2	.0100	.0101	- .0001	- .0100
5	5	0.5	.0625	.0781	- .0156	- .2000
10	10	0.5	.0625	.0694	- .0069	- .1000
20	20	0.5	.0625	.0658	- .0033	- .0500
30	30	0.5	.0625	.0647	- .0022	- .0333
40	40	0.5	.0625	.0641	- .0016	- .0250
50	50	0.5	.0625	.0638	- .0013	- .0200
100	100	0.5	.0625	.0631	- .0006	- .0100
5	5	0.8	.1600	.2000	- .0400	- .2000
10	10	0.8	.1600	.1778	- .0178	- .1000
20	20	0.8	.1600	.1684	- .0084	- .0500
30	30	0.8	.1600	.1655	- .0055	- .0333
40	40	0.8	.1600	.1641	- .0041	- .0250
50	50	0.8	.1600	.1633	- .0033	- .0200
100	100	0.8	.1600	.1616	- .0016	- .0100

Note: All estimates rounded to four digits

* Bias = (Common - Actual)

** Proportional Bias = (Bias ÷ Actual)

Table 2. Estimations of r and f^2 with unbalanced sample sizes

N1	N2	*d*	Common Formula	Actual (Corrected Formula)	Bias *	Proportional Bias **
			Estimating *r* from *d*			
10	10	0.2	.0995	.1048	-.0053	-.0508
11	9	0.2	.0990	.1043	-.0053	-.0508
12	8	0.2	.0975	.1027	-.0052	-.0508
13	7	0.2	.0950	.1001	-.0051	-.0508
14	6	0.2	.0913	.0962	-.0049	-.0509
15	5	0.2	.0863	.0909	-.0047	-.0509
16	4	0.2	.0798	.0840	-.0043	-.0510
17	3	0.2	.0712	.0751	-.0038	-.0510
18	2	0.2	.0599	.0631	-.0032	-.0511
10	10	0.5	.2425	.2548	-.0123	-.0482
11	9	0.5	.2414	.2536	-.0122	-.0483
12	8	0.5	.2379	.2500	-.0121	-.0483
13	7	0.5	.2320	.2438	-.0118	-.0485
14	6	0.5	.2233	.2348	-.0114	-.0487
15	5	0.5	.2116	.2225	-.0109	-.0490
16	4	0.5	.1961	.2063	-.0102	-.0493
17	3	0.5	.1758	.1850	-.0092	-.0497
18	2	0.5	.1483	.1562	-.0078	-.0502
10	10	0.8	.3714	.3885	-.0171	-.0441
11	9	0.8	.3698	.3869	-.0171	-.0441
12	8	0.8	.3649	.3818	-.0169	-.0443
13	7	0.8	.3565	.3732	-.0167	-.0446
14	6	0.8	.3442	.3605	-.0163	-.0451
15	5	0.8	.3273	.3430	-.0157	-.0457
16	4	0.8	.3048	.3196	-.0148	-.0464
17	3	0.8	.2747	.2883	-.0137	-.0473
18	2	0.8	.2334	.2453	-.0119	-.0485
			Estimating f^2 from *d*			
10	10	0.2	.0100	.0111	-.0011	-.1000
11	9	0.2	.0099	.0110	-.0011	-.1000
12	8	0.2	.0096	.0107	-.0011	-.1000
13	7	0.2	.0091	.0101	-.0010	-.1000
14	6	0.2	.0084	.0093	-.0009	-.1000
15	5	0.2	.0075	.0083	-.0008	-.1000
16	4	0.2	.0064	.0071	-.0007	-.1000
17	3	0.2	.0051	.0057	-.0006	-.1000
18	2	0.2	.0036	.0040	-.0004	-.1000
10	10	0.5	.0625	.0694	-.0069	-.1000
11	9	0.5	.0619	.0688	-.0069	-.1000
12	8	0.5	.0600	.0667	-.0067	-.1000
13	7	0.5	.0569	.0632	-.0063	-.1000
14	6	0.5	.0525	.0583	-.0058	-.1000
15	5	0.5	.0469	.0521	-.0052	-.1000
16	4	0.5	.0400	.0444	-.0044	-.1000
17	3	0.5	.0319	.0354	-.0035	-.1000
18	2	0.5	.0225	.0250	-.0025	-.1000
10	10	0.8	.1600	.1778	-.0178	-.1000
11	9	0.8	.1584	.1760	-.0176	-.1000
12	8	0.8	.1536	.1707	-.0171	-.1000
13	7	0.8	.1456	.1618	-.0162	-.1000
14	6	0.8	.1344	.1493	-.0149	-.1000
15	5	0.8	.1200	.1333	-.0133	-.1000
16	4	0.8	.1024	.1138	-.0114	-.1000
17	3	0.8	.0816	.0907	-.0091	-.1000
18	2	0.8	.0576	.0640	-.0064	-.1000

Note: All estimates rounded to four digits

* Bias = (Common - Actual)

** Proportional Bias = (Bias ÷ Actual)

BEST COPY AVAILABLE

References

Wolf, F.M. (1986). "Meta-analysis: Quantitative methods for research synthesis". Sage University Paper series on Quantitative Applications in the Social Sciences, 07-059. Beverly Hills: Sage Pubns.

Rosenthal, R. (1994). Parametric measures of effect size, In H. Cooper & L.V. Hedges, (eds.), The handbook of research synthesis (pp 232-244).

Cohen, J. and Cohen P. (1983). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (2nd Ed.). Hillsdale, NJ: LEA.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: LEA.

Friedman, H. (1968). Magnitude of experimental effect and a table for its rapid estimation. Psychological Bulletin 70:245-251.

Glass, G.V., McGraw, B., and Smith M.L. (1981). Meta-analysis in social research. Sage: Beverly Hills.

Hedges, L.V. and Olkin, I. (1985). Statistical methods for meta-analysis. Academic Press: Orlando.

Rosenthal, R. (1984). Meta-analysis procedures for social research.. Sage: Beverly Hills.